World Energy Production and Consumption: A Visual Study

Aakanksha Kumari | aakankshak.2017@mitb.smu.edu.sg

Amrutha Rajeshwari Yejerla | amruthary.2017@ mitb.smu.edu.sg

Priyanka Sharma | priyankas.2017@mitb.smu.edu.sg

Guided by Professor: Dr. Kam Tin Seong

Abstract—Properly functioning markets for electricity, natural gas, oil and pollution allowances are essential for the rational allocation of resources and cost-effective attainment of environmental goals. Rapidly changing energy dynamics determine the course of our economic development, geopolitics, technological breakthroughs, massive investments and trade flows.With renewable energy moving from niche to mainstream, the fuel mix is shifting. The intent of our project is to visually explore the world in transformation as the global energy system is undergoing major shifts. By using interactive visualization techniques, we propose to aid discovery of long run transitions and short run adjustments by analyzing key factors like production and consumption of various energy resources-both renewable and non-renewable, electricity generation using these resources and the effects of energy usage in terms of Carbon Dioxide emissions.

Index Terms—BP Annual Report, Sunburst Chart, Energy Production and Consumption, Energy Transformation,Non-renewables and Renewables, Geofacets, Interactive Visualisation, Visual Analysis

1 Introduction

Energy growth drives the well-being and prosperity across the globe. Growing demand for energy must be met in a safe and environmentally conscious manner. Rapidly changing energy dynamics determine the course of our economic development, geopolitics, technological breakthroughs, massive investments and trade flows. The main objective of our work is to build a foundational application to enable visual study of the dynamics of energy production and consumption of different countries of the world to identify key trends and dominant players.

Though reports from various organizations such as EIA, IEA, OPEC, BP are available, most of the visualizations are static and do not aid exploration, limiting the scope of further drilling down to the areas of interest of the viewer. As an outcome of our work, we have built an interactive application using RShiny on British Petroleum Statistical Review data released in year 2017, which enables user to easily filter and derive insights from data. With visual analysis, the rich data-set provided by BP reveals patterns that otherwise are hidden in the existing static visualizations.

This paper provides information on the visualization method selection and visualization development efforts for the application for each of the energy dynamics visualization objectives that we set. Introduction is followed by the motivation and objectives of this research. In section 3 a review on previous works in the field is provided. Section 4 describes the dataset and its preparation for visual modelling. Section 5 describes the design framework as well as visualization methodologies chosen and the rationale behind the decisions. Section 6 provides insights we have derived in the process of the development of the application. Future works are stated in section 7 and finally, an installation and user guide details are mentioned in section 8.

2 Objective and Motivation

Data visualization is often an after-thought for many of the practitioners who are collecting and analyzing data. And yet, without clear and compelling communication, analysis will never drive insights and action. Visualization can itself be used effectively in the process of insights discovery. Static visualization and reports do not allow for direct interaction with data, enhanced assimilation of information, quick access to relevant insights and drill-down analysis. By using interactive, visual data analytics techniques we will be bringing the above-mentioned capabilities to the interface. Following are the objectives of our work:

(a) A user-friendly and interactive visualization application interface for data exploration that supports both aggregated and drilled down views and analysis which can be used by personnel in energy departments, policy makers and general public alike.

(b) Interactive visualization to understand which parts of the world today's energy requirements are sourced from.

(c) Regional energy consumption portfolio and trends for each energy type.

(d) Understanding and identifying country clusters as they are set to adopt sustainable energy usage.

3 Previous Work

Visualizations available on the energy outlook reports consists of basic graph types such as bar, line and pie charts which are static in nature and do not facilitate any discovery apart from what they are made to deliver. A sample visualization is presented below:

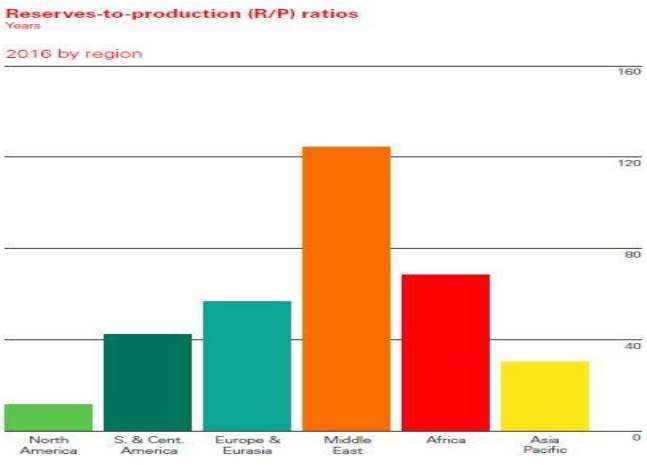


Fig 1: Reserves to Production ratio across regions in BP 2017 Report

There is a lot of scope for improvement in the visualization methodology used and with the open source community contributing

aggressively to the plethora of R packages, the possibilities are endless.

4 Data and its Preparation

For 66 years, the BP Statistical Review of World Energy has provided high-quality objective and globally consistent data on world energy markets. The review is one of the most widely respected and authoritative publications in the field of energy economics, used for reference by the media, academia, world governments and energy companies. A new edition is published every June. For the purpose of this project, we are using the BP Statistical Review of World Energy data released in the year 2017. The dataset consists of energy consumption and production data from around 1965 for 2017 for more than 90 countries spanning across regions Africa, Europe, APAC, CIS, Middle-East , North America, S. & C. America. Energy resources are categorized as Primary Energy(Oil, Coal, Gas, Hydro, Nuclear) and Other Renewable resources(Solar, Wind, Biofuels, Geothermal etc.). We are using data from 2006-2017 for our analysis.

Data provided in the report is present in a single excel workbook containing over 60 tabs and cannot be used as-is. Significant data cleaning and preparation needs to be done in order to make the information usable for the project in R. Data available on multiple sheets is extracted in to R in a long format. Based on the expected input data format of each R visualization package, data is subset, aggregated and reshaped. All the production and consumption parameters for different energy types are measured as (Million Tons of Oil Equivalent) and if this calculation is not available, conversion rates are used to infer values in Mtoe. Data definitions such as the regions that the countries belong to and energy type segregations that are used in the BP report are followed as-is.

5 Design framework and Visualization Methodology

Best analyses show us something that we aren't expecting and when it comes to spotting surprises, "even the best statisticians often set their calculations aside for a while and let their eyes take the lead." (Stephen Few, 2009). Good visual representation practices help in detecting the expected and discovering the unexpected. In order to achieve such a platform, we followed Schneiderman's mantra of overview, zoom, and filter details on demand. For different visualizations we present an overview at the aggregated level (regional level or resource level) and give user the power to drill down and find patterns and insights.

5.1 Energy production distribution

Sunburst is a form of a radial chart which breaks out general categories into subsets to better understand the components that make up or contribute to the whole. This form of visualization is similar to a treemap however it is easier to see multiple layers of data with sunburst, while the treemap is better for comparing categories within the same hierarchical layer. First visual in Fig 2 in the visualization represents the country-wise percentage production across various energy types. Sunburst plot is interactive and can be used to dig deeper to understand world energy production by energy type and region. For example, as revealed in second visual in Figure 2, China produces 13.4% of world's energy from coal.

However, in this visualization deeper slices exaggerate their size, and look visually larger. This type of visualization requires the quantitative comparison of angles, instead of lengths, which is difficult for the human eye.

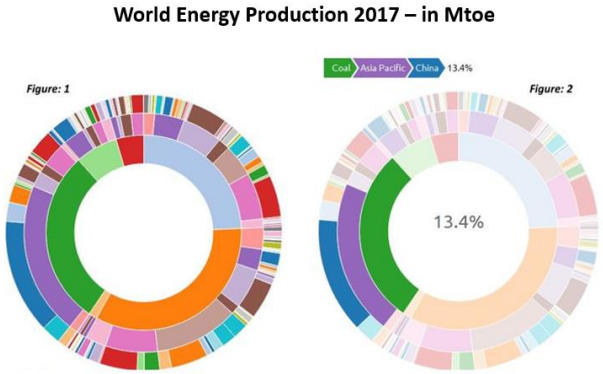


Fig 2: Region-resource type contribution to World energy production 2017

5.2 Primary Energy consumption 2017 across countries

To understand who are the major primary energy consumers of the world and their respective CO2 emissions, we have implemented an interactive world map and a bar graph both in plotly, which are then integrated using shiny as shown in the figure below. World map is being made using leaflet and icon markers are used to indicate OECD and OPEC countries.

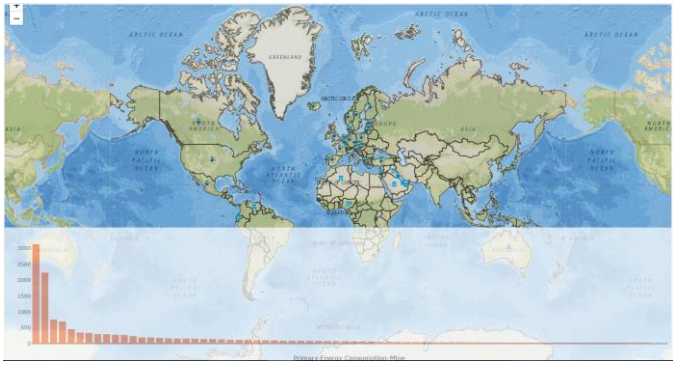


Fig 3 : Primary energy consumption 2017

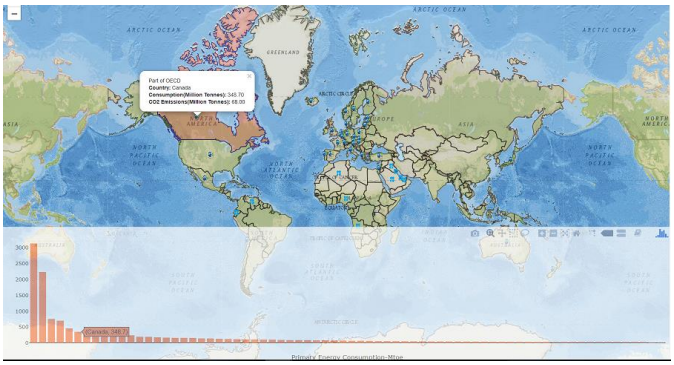


Fig 4: Primary energy consumption 2017- Bar highlighted

Bars are sorted in descending order and when a user clicks on a bar, corresponding country is highlighted in the map and a tool-tip displays the Energy consumption in Mtoe, along with the CO2 emissions and whether the country belongs to OECD/OPAC. Alternatively, the user can click on any country of interest on the

map to get the information on the primary energy consumption and CO2 emissions for the particular country. This is designed to make the exploration interesting and intuitive. In this case, choropleth map will not be an efficient visualization as countries are custom aggregated in the report and for some countries values are not listed.

5.3 Country-wise energy consumption portfolio and energy trends: Geofacets

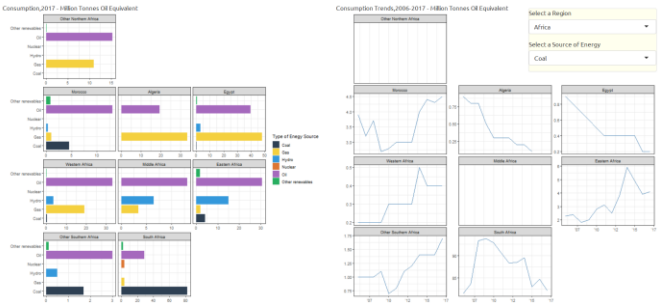


Fig 5: Energy consumption portfolio and trends

Geofacets are used to visualize data for each geographical entity, with the resulting set of visualizations being laid out in a grid that mimics the original geographic topology as closely as possible for the region. In the first figure each Geofacets visualization for a selected region consists of a horizontal bar chart for energy consumption portfolio for all the countries in the region. Second figure in the visualization above shows evolution of energy usage per country over time for each energy type.

Given the custom aggregations of countries in the report as well as the unavailability of data for few countries, Geofacets were the best choice however comparisons across countries form different regions is not direct. Choropleth and Tile maps limit analysis to one variable which is color encoded to aid analysis. Quantifying intensity of color is difficult and choropleths are known to favor large geographic entities over smaller ones. One of the disadvantages of this map is that a Geofacet grid is only meaningful if the person already has an understanding of the underlying original geography. Also, this form of visualization takes more space and representing some of the geographical entities on grid layout can be challenging.

5.4 Identifying country-wise adoption of sustainable energy usage:

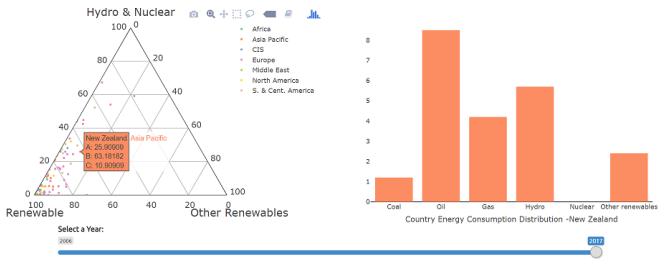


Fig 6: Terenary plot for country positioning

A Terenary plot is used for understanding the relative positioning of countries in terms of usage of Renewable, Hydro & Nuclear energy and Non Renewable energy sources. This plot at a glance helps in identifying countries that source their energy needs from renewables compared to non-renewables etc. Also the bar chart of the absolute measures of usage of each of the energy sources updates with the country selected in ternary chart to aid exploration and discovery. Year slider lets user to visualize the positioning of the country each year. And a slide through years 2006-2017 helps in identifying if a country has taken a serious stand on adopting sustainable energy options.

5.5 Production vs Consumption landscape for each energy type: Interactive Scatter Plot

We have come up with an interactive scatter plot to understand how each country stands in the production vs consumption landscape for each energy type. Total energy production and consumption for each country are calculated and percentage of production and percentage of consumption that the country attributes to each energy type is calculated.

For each of the energy categories namely oil, gas, coal, other renewables (Hydro and Nuclear data is not considered as a pane as production data is not available for these energy types.) countries are plotted on a scatter plot of percentage production vs percentage consumption. The size of the bubble on the scatter plot indicates the absolute amount that the country produces and the color of the bubble indicates the absolute amount that the country consumes of that energy type, as shown in the visualization below.

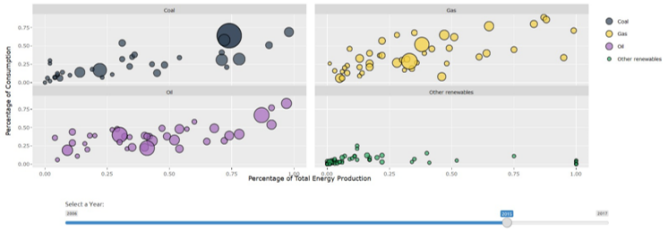


Fig 7: Scatter Plot for Countries – Percentage Production Vs. Consumption

The plot is interactive and the selection of a country/bubble on one pane, highlights the respective positions of the country in other panes aiding insights discovery, as shown below:

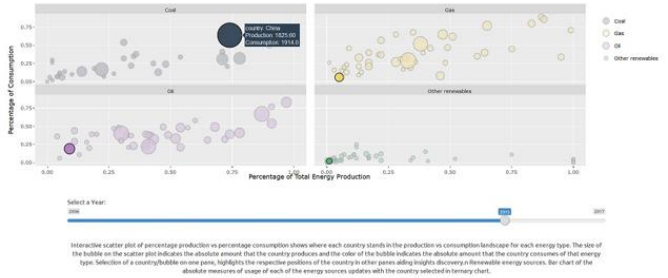
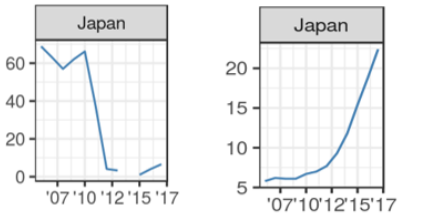


Fig 8: Scatter Plot for Countries – Percentage Production Vs. Consumption - Highlighted

6 Insights and Results

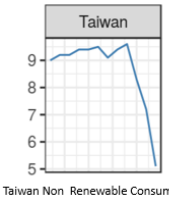
Following are some of the insights from the Visualizations:

Japan is reducing their nuclear power production drastically after 2011 nuclear power plant leak, Fukushima Daiichi due to tsunami and earth quake.



Japan- Nuclear Consumption Japan Non Renewable Consumption

Japan's core focus has shifted from nuclear to other renewables. Last 5 years growth is remarkable, it has replaced about 33% of its nuclear consumption through other renewables.

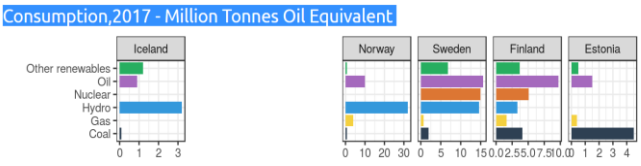


Taiwan Non Renewable Consumption

Taiwan which is at higher risk for tsunami and earthquakes is following similar the same trend and has made significant reduction in their nuclear consumption. In 2016, winning government has stated phasing out nuclear power generation in their agenda.

With the reduction in the prices of Oil, the consumption from 2013 to 2017 has gone up for Oil across all countries. This can be visualized using the time-series pattern in the Geofacet graph for Oil.

For Finland, Iceland & Sweden, it is observed that a major proportion of energy consumption is from other renewable sources of energy.



From the Sunburst plot we can see that China produces 13.4% of the total energy of the world using Coal; Saudi Arabia produces 4.31% of the total energy of the World using Oil

China, US & India are the top 3 countries in the Primary Energy Consumption in 2017

From the Geofacet plots we can see that China has been trying to reduce its Coal consumption after 2012. If we look at Africa, Algeria has been able to reduce its Coal consumption significantly over the years.

The ternary plot shows us that the major energy source for Norway in terms of consumption is Hydro energy, which is a good indication. However, if we look at India, the major energy consumption is for Coal and Oil which are non-renewable sources of energy. US primarily consumes Oil, followed by Gas.

China, over the years has tried to increase the percent contribution from renewable, hydro and nuclear energy sources as opposed to non-renewable energy sources.

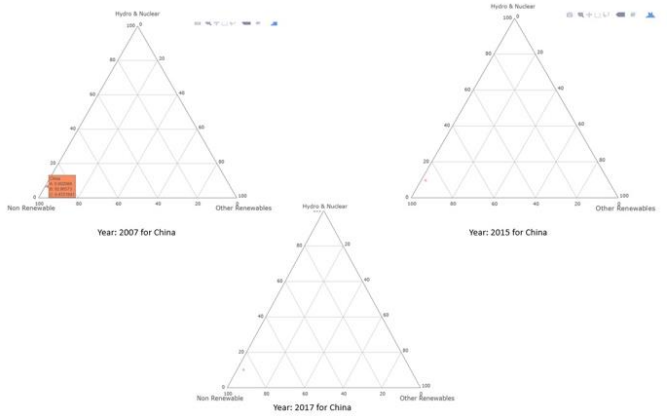


Fig 9: Terenary Plot Overtime - China

7 Conclusion

The original dataset has several other metrics like prices of each energy source, reserves in each country, electricity generation, trade movements etc. These variables also must be included for analysing energy dynamics and the world energy outlook.

Data driven visualization and analysis on energy dataset reveals patterns in production/ consumption of energy across countries with time. The relative contribution of different energy resources in a country's consumptions helps in quantifying the effort that the country puts in moving towards the sustainable energy initiative.

Having said that, the domain knowledge and experience of personnel in energy industry is essential to support or rule out the insights that visual analysis has revealed.

This application is a starting point for interactive analysis of energy data using R visualization packages and shiny application platform. Further, detailed breakdown of renewable sources, energy prices data and trade movements data can also be used to enrich current data and derive additional insights. Adding electricity generation data and power distance would reveal the dynamics of today's world energy landscape.

8 Application Link

Online users would be able to find our application at the following website: https://wiki.smu.edu.sg/1718t3isss608/Group02_Report
Simply click through the tabs to look at the different levels of analysis provided in our application.

9 Acknowledgements

The authors wish to thank Dr Kam Tin Seong for his continual guidance and feedback in applying visualization techniques and packages for the usecase.

10 References

[1] Shneiderman, B. (2005) "The eyes have it: A task by data type taxonomy for information visualization" IEEE Conference on Visual Languages (VL96), pp. 336-343
[2] *About BP:* https://en.wikipedia.org/wiki/BP
[3] *Data:* http://www.bp.com/statisticalreview
[4] http://ryanhafen.com/blog/geofacet
[5] https://hafen.github.io/geofacet/
[6] https://github.com/timelyportfolio/sunburstR
[7] *R Packages Description:* https://cran.r-project.org

11 ANNEX -LIST OF PACKAGES

R Package
dplyr: A Grammar of Data Manipulation. It is a fast, consistent tool for working with data frame like objects, both in memory and out of memory.
tidyr:It's designed specifically for data tidying (not general reshaping or aggregating) and works well with 'dplyr' data pipelines
reshape:Casts a molten data frame into the reshaped or aggregated form you want
readr :The goal of 'readr' is to provide a fast and friendly way to read rectangular data (like 'csv', 'tsv', and 'fwf'). It is designed to flexibly parse many types of data found in the wild, while still cleanly failing when data unexpectedly changes
ggplot:A system for 'declaratively' creating graphics. You provide the data, tell 'ggplot2' how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details
Plotly:Easily translate 'ggplot2' graphs to an interactive web-based version and/or create custom web-based visualizations directly from R
SunburstR:Make interactive 'd3.js' sequence sunburst diagrams in R with the convenience and infrastructure of an 'htmlwidget'.
Crosstalk:Provides building blocks for allowing HTML widgets to communicate with each other, with Shiny or without (i.e. static .html files)
Geofacet:Provides geofaceting functionality for 'ggplot2'. Geofaceting arranges a sequence of plots of data for different geographical entities into a grid that preserves some of the geographical orientation
rgdal:Bindings for the 'Geospatial' Data Abstraction Library
leaflet: Library to create Interactive Web Maps with the JavaScript 'Leaflet'
shiny: Web Application Framework for R
shinythemes: Themes for use with Shiny. Includes several Bootstrap themes
shinydashboard: Create dashboards with 'Shiny'. This package provides a theme on top of 'Shiny', making it easy to create attractive dashboards